# Assessing the number of mean-square derivatives of a Gaussian process

Delphine Blanke [a] , Céline Vial [b,c]

[a] *Université Pierre et Marie Curie-Paris 6, L.S.T.A., 175 rue du Chevaleret, 8ème étage, Bat. A, 75013 Paris, France*

[b] *Université Paris 10, Modal'X, batiment G, 200 avenue de la République, 92000 Nanterre, France*

[c] *Université Pierre et Marie Curie-Paris 6, L.P.M.A., 175 rue du Chevaleret, 4ème étage, Bat. D, 75013 Paris, France*

**Abstract**

We consider a real Gaussian process $X$ with unknown smoothness $r_0 \in \mathbb{N}_0$ where the mean-square derivative $X^{(r_0)}$ is supposed to be Hölder continuous in quadratic mean. First, from the discrete observations $X(t_1), \dots, X(t_n)$, we study reconstruction of $X(t)$, $t \in [0,1]$ with $\widetilde{X}_r(t)$, a piecewise polynomial interpolation of degree $r \geq 1$. We show that the mean-square error of interpolation is a decreasing function of $r$ but becomes stable as soon as $r \geq r_0$. Next, from an interpolation-based empirical criterion, we derive an estimator $\widehat{r}$ of $r_0$ and prove its strong consistency by giving an exponential inequality for $P(\widehat{r} \neq r_0)$. Finally, we prove the strong consistency of $\widetilde{X}_{\widehat{r}}(t)$ with an almost optimal rate.

*Key words:* Gaussian process, Mean-square derivative, Statistical inference, Reconstruction, Piecewise Lagrange interpolation.
*1991 MSC:*
*Running headline:* Assessing the number of derivatives.

*Email addresses:* dblanke@ccr.jussieu.fr (Delphine Blanke),
cvial@u-paris10.fr (Céline Vial).

## 1 Introduction

Let $X = \{X_t,\, t \in [0,1]\}$ be a real Gaussian process whose $r_0$-th ($r_0 \in \mathbb{N}_0$) derivative satisfies a Hölder condition in quadratic mean with exponent $\beta_0 \in [0,1[$ ($r_0$, $\beta_0$ both unknown). In this paper, we propose and study an estimator of $r_0$ based on observations of $X$ at instants $T = \{t_1, \ldots, t_N\} \subset [0,1]$.

As far as we can judge, this question had not yet been raised in the statistical literature of stochastic processes. Nevertheless, this topics arises naturally in several problems of estimation or prediction where processes are supposed to belong to some regularity class depending on $r_0$. More precisely concerning prediction, we refer to works of Cuzick (1977), Lindgren (1979) and Bucklew (1985) where processes have quadratic mean derivatives of specified order. For statistical inference, such an information is required in e.g. the works of Istas and Lang (1997), Blanke and Bosq (1997), Sköld and Hössjer (1999), Blanke and Pumo (2003) and Ditlevsen and Sørensen (2004).

In Section 2, we present our main assumptions as well as various classes of processes fulfilling them. Note that processes satisfying to the so-called Sacks and Ylvisaker (SY) conditions of order $r_0$ are included in our examples.
In Section 3, we present some results about piecewise Lagrange polynomial interpolation. In this context, numerous methods had been proposed and studied under various conditions on processes. Recent works under SY conditions include for example : Müller-Gronbach (1996) (orthogonal projection, optimal designs), Müller-Gronbach and Ritter (1997) (linear interpolation, optimal designs), Müller-Gronbach and Ritter (1998) (linear interpolation, adaptive designs). Under Hölder type conditions, one may cite e.g. works of Seleznjev (1996) (linear interpolation) , Seleznjev (2000) (Hermite interpolation splines, optimal designs), Seleznjev and Buslaev (1998) (best approximation order). For a more detailed survey, we refer to the book by Ritter (2000).
Our work comes within the general framework considered by Plaskota, Ritter and Wasilkowski (2002, 2004) but with a different methodology. We present a slight extension of their results by considering interpolator polynomials $\widetilde{X}_r(t)$ with arbitrary degree $r \geq 1$. We essentially show that the mean-square error of interpolation is a decreasing function of $r$ but becomes stable as soon as $r \geq r_0$. This key point allows us to turn to estimation of $r_0$. In Section 4, we compute an interpolation-based empirical criterion and derive an estimator for $r_0$, denoted by $\widehat{r}$. We show that $\widehat{r}$ is strongly consistent and give an exponential bound for $P(\widehat{r} \neq r_0)$. Finally, these results allow us to establish the strong consistency of $\widetilde{X}_{\widehat{r}}(t)$ with an almost optimal rate. Proofs are postponed to the end of the paper.

## 2 The general framework

### *2.1 Assumptions*

Let $X = \{X_t,\, t \in [0,1]\}$ be a real measurable Gaussian process, defined on the probability space $(\Omega,\, \mathcal{A},\, P)$. We will say that $X \in \mathcal{H}(r_0, \beta_0)$ if it fulfills the assumptions A2.1 and A2.2 stated below.

**Assumption 2.1 (A2.1)**
*(i) $X$ has continuous derivatives in quadratic mean up to order $r_0$, $r_0 \in \mathbb{N}_0$, denoted by $X^{(0)}, \ldots, X^{(r_0)}$,*
*(ii)* $\mathrm{E}\left(X^{(r_0)}(s) - X^{(r_0)}(t)\right)^2 \leq C_1^2 \left|s-t\right|^{2\beta_0}$, $(s,t) \in [0,1]^2$, *with* $\beta_0 \in [0,1[$, $r_0 + \beta_0 > 0$ *and* $C_1 > 0$,
*(iii) the mean function of $X$, $\mu(t) = \mathrm{E}\, X(t)$, has a $(r_0+1)$ bounded derivative for $0 \leq t \leq 1$,*
*(iv)* $\mathrm{E}\left(X(t) - \overline{X}_{a,b}(t)\right)^2 \geq C_2^2 \left(\frac{(b-t)(t-a)}{b-a}\right)^{2(r_0+\beta_0)}$, $t \in ]a,b[$, $0 \leq a < b \leq 1$ *where* $\overline{X}_{a,b}(t) = \mathrm{E}\left(X(t) \big/ X(s),\, s \notin ]a,b[\right)$ *and* $C_2 > 0$ *does not depend on the values $a$ and $b$.*

Let us give some precisions concerning Assumption 2.1. Under A2.1(i),(ii), the process is said to be mean-square hölderian, such conditions give upper bounds for approximation. Recall that A2.1(i) implies in particular that $\mu$ is $r_0$-times continuously differentiable with $\mathrm{E}\, X^{(r)}(t) = \mu^{(r)}(t)$ and that the covariance kernel $K(s,t) = \mathrm{Cov}\,(X(s), X(t))$ is also continuously differentiable with $K^{(r,r)}(s,t) = \mathrm{Cov}\,(X^{(r)}(s), X^{(r)}(t))$, $r = 0, \ldots, r_0$. A2.1(iv) is a more technical condition which is involved in lower bounds for approximation. This condition was first introduced in Plaskota, Ritter and Wasilkowski (2002) where a variety of examples are provided. In the next part, we present and develop these classes of processes.

**Assumption 2.2 (A2.2)**
*(i) On* $[0,1]^2 \setminus \{s=t\}$, $K^{(r_0+1,r_0+1)}(s,t)$ *exists and satisfies* $\left|K^{(r_0+1,r_0+1)}(s,t)\right| \leq C_3$ *with* $C_3 > 0$,
*or*
*(ii) On* $[0,1]^2 \setminus \{s=t\}$, $K^{(r_0+1,r_0+1)}(s,t)$ *exists and satisfies* $\left|K^{(r_0+1,r_0+1)}(s,t)\right| \leq C_3 \left|s-t\right|^{-(2-2\beta_0)}$ *with* $C_3 > 0$.

Note that in the following, we use alternatively A2.2(i) or A2.2(ii) for proving the convergence of our estimator. The first one (together with A2.1(iii)) was introduced by Baxter (1956), in the case $r_0 = 0$, to establish a strong limit theorem for Gaussian processes. His result was next generalized by Gladyshev (1961) under several conditions including A2.2(ii) (with $r_0 = 0$). Let us notice that in both cases, existence of $K^{(r_0+1,r_0+1)}(s,s)$ is not required, in fact it is not wanted since then the process would have $(r_0+1)$ derivatives in quadratic mean. In the sequel, our aim is to estimate the maximal $r$ (denoted by $r_0$) such that one has $X \in \mathcal{H}(r,\beta)$. That's why we have excluded, in condition A2.1(ii), the case $\beta_0 = 1$ to avoid any possible problem of identifiability.

### *2.2 Examples*

We now present and complete the examples given in Plaskota et al. (2002) of Gaussian processes satisfying to our Assumptions A2.1 and A2.2.

**Example 1** *$r_0$-fold integrated fractional Brownian motion.* Let us define $X$ by

$$X(t) = \int_0^t \int_0^{s_{r_0}} \int_0^{s_{r_0-1}} \cdots \int_0^{s_2} W_{\beta_0}(s_1)\,\mathrm{d}s_1\,\mathrm{d}s_2 \cdots \mathrm{d}s_{r_0},$$

in that case $X^{(r_0)} = W_{\beta_0}$, $0 < \beta_0 < 1$ is a fractional standard Brownian motion : $X$ is a zero-mean Gaussian process with covariance function $K^{(r_0,r_0)}(s,t) = \frac{1}{2}(s^{2\beta_0} + t^{2\beta_0} - |s-t|^{2\beta_0})$. It is shown in Plaskota et al. (2002) that $X$ fulfills conditions A2.1(ii)(iv). Moreover, since one has $\mathrm{E}\,(X^{(r_0)}(t+\tau) - \mathrm{E}\,X^{(r_0)}(t))^2 = |\tau|^{2\beta_0}$, condition A2.2(ii) is also satisfied (see Gladyshev (1961)). Finally note that $\beta_0 = 1/2$ yields the $r_0$-fold integrated standard Brownian motion and, in this case, one gets (Plaskota et al. (2002)) : $C_1 = 1$ and $C_2^2 = 1/((2r_0+1)(r_0!)^2)$.

**Example 2** *Sacks-Ylvisaker (SY) conditions.* We take SY conditions as stated in Ritter (2000) p. 68 in the case of a zero-mean process. Let $\Omega_+ = \{(s,t) \in ]0,1[^2 \colon s > t\}$, $\Omega_- = \{(s,t) \in ]0,1[^2 \colon s < t\}$ and let cl $A$ the closure of a set $A$.

(A) $K \in C^{r_0,r_0}([0,1]^2)$, the partial derivatives of $L = K^{(r_0,r_0)}$ up to order two are continuous on $\Omega_+ \cup \Omega_-$ and continuously extendible to cl $\Omega_+$ as well as to cl $\Omega_-$.
(B) $L_-^{(1,0)}(s,s) - L_+^{(1,0)}(s,s) = 1$, $0 \le s \le 1$ where $L_j$ denotes the extension of $L$ to $[0,1]^2$ continuous on cl $\Omega_j$ and on $[0,1]^2 \backslash \mathrm{cl}\,\Omega_j$ with $j \in \{-,+\}$.
(C) $L_+^{(2,0)}(s,\cdot) \in H(L)$ for all $0 \le s \le 1$ and $\sup_{0\le s\le 1} \left\| L_+^{(2,0)}(s,\cdot)\right\|_L < \infty$ where $H(L)$ is a Hilbert space with reproducing kernel $L$ and norm $\|\cdot\|_L$.
(D) If $r_0 \ge 1$, $K^{(r_0,k)}(\cdot,0) = 0$, $k = 0,\ldots,r_0-1$.

First note that for $r_0 \ge 1$, the stationary case is excluded since in this case (D) becomes contradictory with the others conditions. Next, (A) and (B) imply that the process has exactly $r_0$ derivatives in the mean-square sense. Moreover, conditions A2.1(i)(ii) and A2.2(i) are clearly satisfied with $\beta_0 = 1/2$ and $C_1^2 = 2 \sup_{s,t\in\Omega_+\cup\Omega_-} \left|L^{(1,0)}(s,t)\right|$. Let us turn to condition A2.1(iv) : following

Plaskota et al. (2002), it is sufficient to get

$$\sup\{h^2(t)\ : h \in B(K),\ \text{supp}\ h \subseteq [a,b]\} \geq C_2^2\Big(\frac{(b-t)(t-a)}{b-a}\Big)^{2(r_0+\beta_0)}$$

where $B(K)$ denotes the unit ball in $H(K)$. Next applying results of Proposition IV.8 of Ritter (2000) p. 78 we successively obtain for $0 \leq a < b \leq 1$ and $t \in ]a,b[$ :

$$\begin{aligned}\sup\Big\{h^2(t) : h \in B(K),\ \text{supp}\ h \subseteq [a,b]\Big\} &\geq \sup\Big\{\|h\|_K^{-2} \cdot h^2(t) : h \in H(K) \cap \Delta_{a,b}\Big\}\\ &\geq \sup\Big\{\|h\|_K^{-2} \cdot h^2(t)\ :\ h \in W_2^{r_0+1}([0,1]) \cap \Delta_{a,b}\Big\}\\ &\geq \sup\Big\{\|h\|_K^{-2}\left\|h^{(r_0+1)}\right\|_2^2 \cdot h^2(t)\ :\ h \in B_{r_0} \cap \Delta_{a,b}\Big\}\\ &\geq (1+c)^{-2} \cdot \frac{1}{(2r_0+1)\cdot(r_0!)^2}\Big(\frac{(b-t)(t-a)}{b-a}\Big)^{2r_0+1}, \qquad (2.1)\end{aligned}$$

for a positive constant $c$ depending only on $K$ and $B_{r_0}$ corresponding to the unit ball in $W_2^{r_0+1}([0,1])$ where

$$W_2^{r_0+1}([0,1]) = \Big\{g \in C^{r_0}([0,1])\ :\ g^{(r_0)}\ \text{abs. cont.}\ , g^{(r_0+1)} \in L^2([0,1])\Big\}$$

and $\Delta_{a,b} = \Big\{h\ :\ h^{(k)}(a) = h^{(k)}(b) = 0,\ k = 0,\ldots,r_0,\ \text{supp}\ h = [a,b], h(t) \neq 0 \text{ for all } t \in ]a,b[\Big\}$.

Note that the solution of the extremal problem given in (2.1) had be obtained by Speckman (1979) (see also Ritter (2000), p. 94). Thereby, the condition A2.1(iv) is checked with $\beta_0 = 1/2$ and $C_2 = (2r_0+1)^{-1} \cdot ((1+c)r_0!)^{-2}$.

**Example 3** *Stationary processes with spectral density $\varphi$.* Suppose that $\varphi$ satisfies both for $u$ large enough, $\varphi(u) \leq c_1 |u|^{-2\gamma}$ with $c_1 > 0$, $\gamma > 1/2$, and for every real $u$, $\varphi(u) \geq c_0(1+u^2)^{-\gamma}$ with $c_0 > 0$, then results of Plaskota et al. (2002) imply that for $\gamma - \frac{1}{2} \notin \mathbb{N}$ the conditions A2.1(ii)(iv) are fulfilled with $r_0 = \left[\gamma - \frac{1}{2}\right]$ and $\beta_0 = \gamma - \frac{1}{2} - r_0$. Next, we strengthen the first condition by $\left|\varphi(u) - c_1 |u|^{-2\gamma}\right| \leq c_2 |u|^{-2(\gamma+1)}$ with $c_2 > 0$, $\gamma > 1/2$ and $u$ large enough. In this case, from $K^{(r_0,r_0)}(s,t) = \int_{-\infty}^{\infty} u^{2r_0} e^{\imath(s-t)u}\varphi(u)\,\mathrm{d}u$ and by adapting the proof of Gladyshev (1961) to the case $r_0 \geq 1$, we obtain the required condition A2.2(ii). For instance, one has $K(s,t) = (2\theta)^{-1}\exp(-\theta|s-t|)$ and $\varphi(u) = \pi^{-2}(\theta^2+u^2)^{-1}$ for an Ornstein-Uhlenbeck (O.U.) process, which implies in turn that $\gamma = 1$, $r_0 = 0$ and $\beta_0 = 1/2$.

**Example 4** *$r_0$-fold integrated stationary processes.* Let $Y = \{Y_t,\ t \in [0,1]\}$ be a zero-mean stationary Gaussian process with covariance $\rho_0(|t-s|)$. Lasinger (1993) gives conditions under which stationarity is preserved by $r_0$-fold integration. Namely, one may define recursive stationary processes $X^{(r_0-i)} =$

$\left\{X_t^{(r_0-i)},\, t\in[0,1]\right\}$ by $X_t^{(r_0-i)} = X_0^{(r_0-i)} + \int_0^t X_s^{(r_0-i+1)}\,\mathrm{d}s$, for $i=1,\dots,r_0$. Here one has $X^{(r_0)} := Y$, $X_0^{(r_0-i)}$ are Gaussian random variables with positive variance $\sigma_i^2$ and $\rho_i(t) := \sigma_i^2 - \int_0^t\int_0^u \rho_{i-1}(v)\,\mathrm{d}v\,\mathrm{d}u$ represents the covariance function of $X^{(r_0-i)}$. It is shown in Lasinger (1993) that such construction is possible when the covariance is either linear : $\rho_0(t) = 1-\lambda\,|t|$ $(0<\lambda<2)$ or exponential : $\rho_0(t) = (2\theta)^{-1}\exp(-\theta\,|t|)$ $(\theta>0)$ as soon as one choose $\sigma_i^2 > \|\rho_i'\|^2_{\rho_{i-1}}$, $i=1,\dots,r_0$ in each integration step. Moreover from Theorem 4 of Lasinger (1993), one can evaluate $\|h\|_K$ (here $K=\rho_{r_0}$) for $h\in H(K)\cap\Delta_{a,b}$ (see Example 2) in the linear case : $\|h\|_K^2 = \left\|h^{(r_0)}\right\|^2_{\rho_0} = (2\lambda)^{-1}\int_0^1 h^{(r_0+1)}(t)^2\,\mathrm{d}t$ and in the exponential one : $\|h\|_K^2 = \left\|h^{(r_0)}\right\|^2_{\rho_0} = \int_0^1 h^{(r_0+1)}(t)^2\,\mathrm{d}t + \theta^2\int_0^1 h^{(r_0)}(t)^2\,\mathrm{d}t$. Then same methodology as in Example 2 (using Lemma IV.4 of Ritter (2000) p. 73) yields the required condition A2.1(iv) with $\beta_0 = 1/2$ in both cases. Note that conditions A2.1(ii), A2.2 are stated on $K^{(r_0,r_0)}(s,t) = \rho_0(|s-t|)$ and consequently are easily checked. Finally, the linear case occurs for example when $X^{(r_0)}(t) = W(t+1)-W(t)$ whereas the exponential one corresponds to the O.U. process (see Example 3).

**Example 5** *Non centered case.* It is easy to see that if $Z_t = X_t + \mu(t)$ where $X = (X_t,\, t\in[0,1])$ is a zero-mean Gaussian process satisfying to conditions A2.1(i)(ii)(iv) and A2.2 then $Z = (Z_t,\, t\in[0,1])$ fulfills Assumption A2.1, A2.2 as soon as $\mu$ checks A2.1(iii).

## 3 Results on piecewise Lagrange interpolation

Suppose that $X$ is observed at the $(nr+1)$ equally spaced instants $s_{\ell,r,n} := s_{\ell,r} = \frac{\ell}{r}\delta_n$ with $\ell = 0,\dots,nr$, $\delta_n = 1/n$ and for some $r\in\mathbb{N}$. Next we will use the knots $t_{i,k,n} := t_{i,k} = k\delta_n + \frac{i}{r}\delta_n$ with $i=0,\dots,r$ and $k=0,\dots,n-1$. Note that some of the $t_{i,k}$ are equal, in particular $t_{r,k} = t_{0,k+1}$. Actually, one has $t_{i,k} = s_{kr+i,r}$ and for $[\ell/r]\neq \ell/r$, $s_{\ell,r} = t_{\ell-r[\frac{\ell}{r}],[\frac{\ell}{r}]}$ whereas if $[\ell/r] = \ell/r$, $s_{\ell,r} = t_{0,[\frac{\ell}{r}]} = t_{r,[\frac{\ell}{r}]-1}$ where $[.]$ denotes the integer part. The approximation process is then defined by

$$\widetilde{X}_r(t) = \sum_{i=0}^{r} L_{i,k}(t)X(t_{i,k}), \tag{3.1}$$

for $t\in I_k := [k\delta_n, (k+1)\delta_n]$ and where $L_{i,k}(t)$ is the Lagrange interpolator polynomial given by

$$L_{i,k}(t) = \prod_{\substack{j=0\\ j\neq i}}^{r} \frac{(t-t_{j,k})}{(t_{i,k}-t_{j,k})}. \tag{3.2}$$

Recall that for $t \in I_k$, $\widetilde{X}_r(t)$ corresponds, by definition, to the unique polynomial of degree (at most) $r$ which interpolates $X(t)$ at the $(r+1)$ knots $t_{i,k}$, $i = 0, \ldots, r$. Using these polynomials presents several advantages : they are easy to build and to implement, they give sharp upper bounds for approximation (see Proposition 3.1 and its following remarks). Moreover conversely to Hermite polynomials, they do not require observation of derivatives of the process $X$. Comparing to splines, one classical drawback is that the derivatives $\widetilde{X}_r^{(\ell)}$, $\ell = 1, \ldots, r$, are typically discontinuous at the knots $t_{i,j}$. Anyway all our conditions are stated in a mean-square sense so pathwise irregularities are of no importance in our context.

### *3.1 A preliminary result*

For further results, we have to express the function $\mathcal{C}_r(s,t) := \mathrm{E}\Big(X(t) - \widetilde{X}_r(t)\Big)\Big(X(s) - \widetilde{X}_r(s)\Big)$ with $r \geq 1$, in terms of the covariance $K(s,t)$ and its partial derivatives for $(t,s) \in [k\delta_n, (k+1)\delta_n] \times [\ell\delta_n, (\ell+1)\delta_n]$.

**Lemma 3.1**
*Suppose that condition A2.1(i) holds and let $r^* = \min(r_0, r)$ be such that $r^* \geq 1$. If $t = k\delta_n + \theta_1\delta_n$, $s = \ell\delta_n + \theta_2\delta_n$ with $0 \leq \theta_1, \theta_2 \leq 1$, one obtains,*

$$\mathcal{C}_r(s,t) = \sum_{i,j=0}^{r} \frac{L_{i,k}(t)L_{j,\ell}(s)(ij\delta_n^2)^{r^*}}{(r^{r^*}(r^*-1)!)^2} \cdot \iint\limits_{[0,1]^2} \Big((1-v)(1-w)\Big)^{r^*-1} \mathcal{R}_{\delta_n,i,j}(v,w)\,\mathrm{d}v\,\mathrm{d}w \tag{3.3}$$

*where*

$$\begin{aligned}\mathcal{R}_{\delta_n,i,j}(v,w) = {} & K^{(r^*,r^*)}(k\delta_n + \theta_1\delta_n v, \ell\delta_n + \theta_2\delta_n w) - K^{(r^*,r^*)}(k\delta_n + \theta_1\delta_n v, \ell\delta_n + \frac{j\delta_n}{r}w) \\ & - K^{(r^*,r^*)}(k\delta_n + \frac{i\delta_n}{r}v, \ell\delta_n + \theta_2\delta_n w) + K^{(r^*,r^*)}(k\delta_n + \frac{i\delta_n}{r}v, \ell\delta_n + \frac{j\delta_n}{r}w) \\ & + \mu^{(r^*)}(k\delta_n + \theta_1\delta_n v)\mu^{(r^*)}(\ell\delta_n + \theta_2\delta_n w) - \mu^{(r^*)}(k\delta_n + \theta_1\delta_n v)\mu^{(r^*)}(\ell\delta_n + \frac{j\delta_n}{r}w) \\ & - \mu^{(r^*)}(k\delta_n + \frac{i\delta_n}{r}v)\mu^{(r^*)}(\ell\delta_n + \theta_2\delta_n w) + \mu^{(r^*)}(k\delta_n + \frac{i\delta_n}{r}v)\mu^{(r^*)}(\ell\delta_n + \frac{j\delta_n}{r}w).\end{aligned} \tag{3.4}$$

**Remarks :** (i) In the case $r_0 = 0$, one may easily obtain

$$\mathcal{C}_r(s,t) = \sum_{i,j=0}^{r} L_{i,k}(t)L_{j,\ell}(s)\cdot\Big\{K(t,s) - K(t,\ell\delta_n + \frac{j\delta_n}{r}) - K(k\delta_n + \frac{i\delta_n}{r}, s)$$
$$+ K(k\delta_n + \frac{i\delta_n}{r}, \ell\delta_n + \frac{j\delta_n}{r}) + \mu(t)\mu(s) - \mu(t)\mu(\ell\delta_n + \frac{j\delta_n}{r})$$
$$- \mu(k\delta_n + \frac{i\delta_n}{r})\mu(s) + \mu(k\delta_n + \frac{i\delta_n}{r})\mu(\ell\delta_n + \frac{j\delta_n}{r})\Big\}. \quad (3.5)$$

In fact, note that for $(t,s) \in [k\delta_n, (k+1)\delta_n] \times [\ell\delta_n, (\ell+1)\delta_n]$ :

$$\mathcal{C}_r(s,t) = \mathrm{E}\,X(t)X(s) + \sum_{i,j=0}^{r} L_{i,k}(t)L_{j,\ell}(s)\mathrm{E}\,X(\frac{i\delta_n}{r} + k\delta_n)X(\frac{j\delta_n}{r} + \ell\delta_n)$$
$$- \sum_{j=0}^{r} L_{j,\ell}(s)\mathrm{E}\,X(t)X(\frac{j\delta_n}{r} + \ell\delta_n) - \sum_{i=0}^{r} L_{i,k}(t)\mathrm{E}\,X(\frac{i\delta_n}{r} + k\delta_n)X_s$$

and the result follows since uniqueness of Lagrange polynomials implies in particular $\sum_{i=0}^{r} L_{i,k}(t) = \sum_{j=0}^{r} L_{j,\ell}(s) = 1$.

(ii) In the stationary case where $\mu(t) \equiv \mu$, the term $\mathcal{R}_{\delta_n,i,j}$, given by (3.4), simply reduces to

$$\mathcal{R}_{\delta_n,i,j}(v,w) = K^{(r^*,r^*)}(k\delta_n + \theta_1\delta_n v, \ell\delta_n + \theta_2\delta_n w) - K^{(r^*,r^*)}(k\delta_n + \theta_1\delta_n v, \ell\delta_n + \frac{j\delta_n}{r}w)$$
$$- K^{(r^*,r^*)}(k\delta_n + \frac{i\delta_n}{r}v, \ell\delta_n + \theta_2\delta_n w) + K^{(r^*,r^*)}(k\delta_n + \frac{i\delta_n}{r}v, \ell\delta_n + \frac{j\delta_n}{r}w). \quad (3.6)$$

### *3.2 Upper and lower bounds for the error of interpolation*

Using properties of reproducing kernel Hilbert spaces, Plaskota et al. (2002) give error estimates for piecewise Lagrange interpolation of order $r \geq r_0$. Our lemma 3.1 allows us to generalize their result to any order $r \geq 1$. Recall that here and in all the following, $\delta_n = n^{-1}$.

**Proposition 3.1**
*(a) Under conditions A2.1(i)(ii), we obtain for any $r \geq 1$,*

$$\sup_{t\in[0,1]} \mathrm{E}\left(X(t) - \widetilde{X}_r(t)\right)^2 \leq A_1^2(r)\delta_n^{2(r^*+\beta^*)}, \quad (3.7)$$

*with* $r^* = \min(r, r_0)$, $\beta^* = \begin{cases} 1 & \text{if } r < r_0 \\ \beta_0 & \text{if } r \geq r_0, \end{cases}$

*and* $A_1^2(r) = \begin{cases} A_1^2(r, \sup_{s,t\in[0,1]^2} \left|K^{(r+1,r+1)}(s,t)\right|) & \text{if } r < r_0 \\ A_1^2(r, C_1) & \text{if } r \geq r_0. \end{cases}$

*(b) Under condition A2.1(iii), one has for any* $r \geq 1$*,*

$$\sup_{t\in[0,1]} \left|\mathrm{E}\left(X(t) - \widetilde{X}_r(t)\right)\right| \leq A_2(r, \|\mu^{(r^*+1)}\|_\infty)\delta_n^{r^*+1}. \tag{3.8}$$

**Remarks**

- As our upper bounds only involve the covariance kernel of the process $X$, they hold also for non Gaussian processes satisfying to assumptions A2.1.
- In the stationary case where $\mu(t) \equiv \mu$, it is easy to see that, the approximation is unbiased : $\mathrm{E}\left(X(t) - \widetilde{X}_r(t)\right) = 0$, $t \in [0,1]$.
- For $r \geq r_0$, the rate of approximation $n^{-(r_0+\beta_0)}$ appearing in (3.7) is optimal in several senses for the Hölder class $\mathcal{H}(r_0, \beta_0)$ : see Seleznjev and Buslaev (1998).

Now, let us turn to a lower bound of approximation.

**Proposition 3.2**
*Let* $\dot{s}_{k,r}$ *be the middle of* $[s_{k,r},\ s_{k+1,r}]$ *(with* $s_{k,r} = \frac{k}{r}\delta_n$*,* $k = 0, \ldots, nr$*), then the condition A2.1(iv) implies for any* $r \geq 1$ *:*

$$\mathrm{E}\left(X(\dot{s}_{k,r}) - \widetilde{X}_r(\dot{s}_{k,r})\right)^2 \geq A_3^2(r, C_2)\, \delta_n^{2(r_0+\beta_0)}.$$

Note that a similar result for $\int_0^1 \mathrm{E}\left(X(t) - \mathcal{A}X(t)\right)^2 \mathrm{d}t$ is obtained in Plaskota et al. (2002) for any algorithm $\mathcal{A}$ using the knots $s_{k,r}$.

## 4 Estimation of the parameter $r_0$

### *4.1 The estimator*

Let us assume that $(X_t,\ t \in [0,1])$ is observed at points $s_{k,r} = \frac{kr}{n}$ (with $r = 1, \ldots, p_n$ and $k = 0, \ldots, nr$) and at points $\dot{s}_{k,r} = \frac{1}{2}(s_{k,r} + s_{k+1,r})$ (with $r = 1, \ldots, p_n$ and $k = 0, \ldots, nr - 1$), which means that the process is observed at $N := np_n(p_n+1)+p_n$ points. In the following we will choose $p_n = \log_a(n) := \underbrace{\ln\ln\cdots\ln}_{a \text{ times}} n$ with $a \geq 2$, chosen arbitrary large.

We define the estimator of the parameter $r_0$ as

$$\widehat{r} = \min\Big\{r \in \{1, \ldots, p_n\} : \frac{1}{nr}\sum_{j=0}^{nr-1}(X(\dot{s}_{j,r}) - \widetilde{X}_r(\dot{s}_{j,r}))^2 \geq n^{-2r}h_n\Big\} - 1$$

and in case of having an empty set, we will set $\widehat{r}$ to an arbitrary value $l_0$.

The threshold $h_n$ is chosen such that $h_n n^{2\beta_0-2} \to 0$ and $h_n n^{2\beta_0} \to \infty$, for all $\beta_0 \in [0, 1[$. For example, an omnibus choice is e.g. $h_n = \ln n$.

**Remark 4.1** *We may notice that as the sequence $p_n$ tends to infinity with $n$, there exists an integer $n_0$ such that $r_0 \in \{0, \ldots, p_n\}$ for all $n \geq n_0$. Furthermore if an upper bound $B$ is known for $r_0$, one can choose $p_n = B + 1$.*

### *4.2 Consistency of the estimator*

We now present the main result of our paper, namely an exponential upper bound for the probability of the event $\{\widehat{r} \neq r_0\}$.

**Theorem 4.1**
*(i) Let us assume the hypothesis A2.1 and A2.2(i),*

$$P(\widehat{r} \neq r_0) = \mathcal{O}\Big(\exp\Big(-D_1 \min\Big(n, \delta_n^{-(2-2\beta_0)}\Big)\Big)\Big);$$

*where $D_1$ is a positive constant.*

*(ii) Let us assume the hypothesis A2.1 and A2.2(ii),*

$$P(\widehat{r} \neq r_0) = \mathcal{O}\Big(\exp\Big(-D_2 \min\Big(\frac{n}{\ln n}, \delta_n^{-(2-2\beta_0)}\Big)\Big)\Big)$$

*where $D_2$ is a positive constant.*

**Remarks**

- Let us note that more explicit bounds are established along the proof, see the relations (5.6)-(5.7).
- In both cases, the rate of convergence is exponential and depends on $\beta_0$ : as expected, the case $\beta_0 = 1$ turns to be degenerate.

Since $N = np_n(p_n + 1) + p_n$, we obtain the following immediate corollary.

**Corollary 4.1** *Under Assumptions A2.1 and A2.2, one has*

*(i) if* $\beta_0 \in [0, \frac{1}{2}]$,

$$P(\widehat{r} \neq r_0) = \mathcal{O}\Big(\exp\Big(-D_3 \frac{N}{(\ln N)(\log_a N)^2}\Big)\Big),$$

*(ii) if* $\beta_0 \in ]\frac{1}{2}, 1[$,

$$P(\widehat{r} \neq r_0) = \mathcal{O}\Big(\exp\Big(-D_4 \Big(\frac{N}{(\log_a N)^2}\Big)^{2-2\beta_0}\Big)\Big),$$

*where $D_3$ and $D_4$ are positive constants.*

To establish Theorem 4.1, we need the following proposition concerning behavior of $\mathcal{E}_r(k, \ell)$ defined by

$$\mathcal{E}_r(k, \ell) = E\left[\Big(X(\dot{s}_{k,r}) - \widetilde{X}_r(\dot{s}_{k,r})\Big)\Big(X(\dot{s}_{\ell,r}) - \widetilde{X}_r(\dot{s}_{\ell,r})\Big)\right],$$

for $k, \ell \in \{0, \ldots, nr-1\}$ and $r = 1, \ldots, r_0 + 1$.

**Proposition 4.1** *Let us assume the hypothesis A2.1 with $\mu(t) \equiv \mu$, then*

*(a) if $r = 1, \ldots, r_0 - 1$ (in the case, $r_0 \geq 2$) :*

$$\max_{0 \leq k \leq nr-1} \frac{\delta_n^{-2r}}{n} \sum_{\ell=0}^{nr-1} |\mathcal{E}_r(k, \ell)| \leq C_1 \delta_n^2. \tag{4.1}$$

*(b) if $r = r_0$ (in the case, $r_0 \geq 1$) or $r = r_0 + 1$, condition A2.2(i) yields*

$$\max_{0 \leq k \leq nr-1} \frac{\delta_n^{-2(r_0+\beta_0)}}{n} \sum_{\ell=0}^{nr-1} |\mathcal{E}_r(k, \ell)| \leq C_2 \max\left(\delta_n^{2(1-\beta_0)}, \frac{1}{n}\right), \tag{4.2}$$

*whereas under A2.2(ii),*

$$\max_{0 \leq k \leq nr-1} \frac{\delta_n^{-2(r_0+\beta_0)}}{n} \sum_{\ell=0}^{nr-1} |\mathcal{E}_r(k, \ell)| \leq C_3 \max\left(\delta_n^{2(1-\beta_0)}, \frac{\ln n}{n}\right), \tag{4.3}$$

Since in our case $\delta_n = n^{-1}$, it follows that, if $r = r_0, r_0 + 1$, one gets an upper bound not exceeding $\ln n/n$ for $\beta_0 \leq 1/2$ and $n^{-2(1-\beta_0)}$ if $\beta_0 > 1/2$.

Finally, Theorem 4.1 allows us to study the pointwise almost sure convergence of $\widetilde{X}_{\widetilde{r}}(t)$ toward $X(t)$, where we have set $\widetilde{r} = \max(\widehat{r}, 1)$.

**Theorem 4.2** *Under Assumptions A2.1 and A2.2,*

*(i)* $\widehat{r} = r_0$ *almost surely for* $N$ *large enough,*

*(ii) for all* $t \in [0,1]$, *one gets for all positive* $\varepsilon$,

$$\frac{N^{(r_0+\beta_0)}}{(\ln N)^{1/2+\varepsilon}} \left| X(t) - \widetilde{X}_{\widetilde{r}}(t) \right| \xrightarrow[N\to\infty]{a.s.} 0.$$

Comparing this last result to the upper bound obtained in Proposition 3.1(a), we see that our rate is not far from the optimal rate of approximation of $X(t)$ by $\widetilde{X}_{r_0}(t)$ (see Seleznjev and Buslaev, 1998).

## 5 Proofs

### *5.1 Proof of Lemma 3.1*

(a) Suppose first that $\mu(t) \equiv \mu$. Recall that the condition A2.1(i) holds iff $K \in C^{r_0,r_0}([0,1]^2)$, in other words, $K$ has continuous partial derivatives $K^{(p,m)}$ on $[0,1]^2$ for all integers $p, m \le r_0$. From (3.5), we apply a Taylor series expansion (with integral remainder) of order $r^* = \min(r, r_0)$ :

$$K(t,s) = \sum_{p=0}^{r^*-1} \frac{(\theta_1\delta_n)^p}{p!} K^{(p,0)}(k\delta_n, s) + \int_0^1 \frac{(\theta_1\delta_n)^{r^*}}{(r^*-1)!}(1-v)^{r^*-1} K^{(r^*,0)}(k\delta_n + \theta_1\delta_n v, s)\,\mathrm{d}v,$$

but one has also,

$$\begin{aligned} K^{(p,0)}(k\delta_n, s) = &\sum_{m=0}^{r^*-1} \frac{(\theta_2\delta_n)^m}{m!} K^{(p,m)}(k\delta_n, \ell\delta_n) \\ &+ \int_0^1 \frac{(\theta_2\delta_n)^{r^*}}{(r^*-1)!}(1-w)^{r^*-1} K^{(p,r^*)}(k\delta_n, \ell\delta_n + \theta_2\delta_n w)\,\mathrm{d}w \end{aligned}$$

and

$$\begin{aligned} K^{(r^*,0)}(k\delta_n + \theta_1\delta_n v, s) = &\sum_{m=0}^{r^*-1} \frac{(\theta_2\delta_n)^m}{m!} K^{(r^*,m)}(k\delta_n + \theta_1\delta_n, \ell\delta_n) \\ &+ \int_0^1 \frac{(\theta_2\delta_n)^{r^*}}{(r^*-1)!}(1-w)^{r^*-1} K^{(r^*,r^*)}(k\delta_n + \theta_1\delta_n v, \ell\delta_n + \theta_2\delta_n w)\,\mathrm{d}w \end{aligned}$$

Now similar expansions hold for $K(t, \ell\delta_n + \frac{j\delta_n}{r})$, $K(k\delta_n + \frac{i\delta_n}{r}, s)$ and $K(k\delta_n + \frac{i\delta_n}{r}, \ell\delta_n + \frac{j\delta_n}{r})$ (with $\theta_1\delta_n$, $\theta_2\delta_n$ respectively replaced by $i\delta_n/r$, $j\delta_n/r$). This yields : $\mathcal{C}_r(t,s) = A_1 + A_2 + A_3 + A_4$ where the $A_i$ are respectively defined by

$$A_1 = \sum_{p,m=0}^{r^*-1} \frac{K^{(p,m)}(k\delta_n, \ell\delta_n)}{p!m!} \sum_{i,j=0}^{r} L_{i,k}(t)L_{j,\ell}(s)\Big\{(\theta_1\delta_n)^p(\theta_2\delta_n)^m - (\theta_1\delta_n)^p(j\delta_n/r)^m$$
$$- (i\delta_n/r)^p(\theta_2\delta_n)^m + (i\delta_n/r)^p(j\delta_n/r)^m\Big\},$$

$$A_2 = \sum_{p=0}^{r^*-1}\sum_{i,j=0}^{r} \frac{L_{i,k}(t)L_{j,\ell}(s)}{(r^*-1)!p!} \cdot \int_0^1 (1-w)^{r^*-1}\Big\{(\theta_2\delta_n)^{r^*}K^{(p,r^*)}(k\delta_n, \ell\delta_n + \theta_2\delta_n w)$$
$$- (j\delta_n/r)^{r^*}K^{(p,r^*)}(k\delta_n, \ell\delta_n + \frac{j\delta_n}{r})\Big\}\,\mathrm{d}w \cdot \Big\{(\theta_1\delta_n)^p - (i\delta_n/r)^p\Big\},$$

$$A_3 = \sum_{m=0}^{r^*-1}\sum_{i,j=0}^{r} \frac{L_{i,k}(t)L_{j,\ell}(s)}{(r^*-1)!m!} \cdot \int_0^1 (1-v)^{r^*-1}\Big\{(\theta_1\delta_n)^{r^*}K^{(r^*,m)}(k\delta_n + \theta_1\delta_n v, \ell\delta_n)$$
$$- (i\delta_n/r)^{r^*}K^{(r^*,m)}(k\delta_n + \frac{i\delta_n}{r}, \ell\delta_n)\Big\}\,\mathrm{d}v \cdot \Big\{(\theta_2\delta_n)^m - (j\delta_n/r)^m\Big\},$$

$$A_4 = \sum_{i,j=0}^{r} \frac{L_{i,k}(t)L_{j,\ell}(s)}{(r^*-1)!^2} \cdot \iint_{[0,1]^2} \Big\{(\theta_1\delta_n)^{r^*}(\theta_2\delta_n)^{r^*}K^{(r^*,r^*)}(k\delta_n + \theta_1\delta_n v, \ell\delta_n + \theta_2\delta_n w)$$
$$- (\theta_1\delta_n)^{r^*}(j\delta_n/r)^{r^*}K^{(r^*,r^*)}(k\delta_n + \theta_1\delta_n v, \ell\delta_n + \frac{j\delta_n}{r}w)$$
$$- (i\delta_n/r)^{r^*}(\theta_2\delta_n)^{r^*}K^{(r^*,r^*)}(k\delta_n + \frac{i\delta_n}{r}v, \ell\delta_n + \theta_2\delta_n w)$$
$$+(i\delta_n/r)^{r^*}(j\delta_n/r)^{r^*}K^{(r^*,r^*)}(k\delta_n + \frac{i\delta_n}{r}v, \ell\delta_n + \frac{j\delta_n}{r}w)\Big\}\Big((1-v)(1-w)\Big)^{r^*-1}\,\mathrm{d}v\,\mathrm{d}w.$$

Now, one may easily obtain $A_1 = A_2 = A_3 = 0$. Consider for example the term $\sum_{j=0}^{r} L_{j,\ell}(s)(j\delta_n/r)^m$ : it corresponds to the unique polynomial of degree at most $r$ which interpolates the function taking values $(j\delta_n/r)^m$ at points $\ell\delta_n + \frac{j}{r}\delta_n$ with $j = 0, \ldots, r$. Then $(x - \ell\delta_n)^m$ (with $m = 0, \ldots, r^* - 1 \le r - 1$) is adequate and $\sum_{j=0}^{r} L_{j,\ell}(s)(j\delta_n/r)^m = (s - \ell\delta_n)^m = (\theta_2\delta_n)^m$. By the same way, $\sum_{i=0}^{r} L_{i,k}(t)(i\delta_n/r)^p = (\theta_1\delta_n)^p$. This implies in turn that summations over $(i,j)$ in $A_1$, over $i$ in $A_2$ and $j$ in $A_3$ respectively vanish.

The same trick (since $r^* \le r$ and $\sum_{i=0}^{r} L_{i,k}(t) = \sum_{j=0}^{r} L_{j,\ell}(s) = 1$) leads to

$$A_4 = \sum_{i,j=0}^{r} \frac{L_{i,k}(t)L_{j,\ell}(s)(ij\delta_n^2)^{r^*}}{r^{2r^*}(r^*-1)!^2} \cdot \iint_{[0,1]^2} \Big((1-v)(1-w)\Big)^{r^*-1}\mathcal{R}_{\delta_n,i,j}(v,w)\,\mathrm{d}v\,\mathrm{d}w$$

with $\mathcal{R}_{\delta_n,i,j}$ given by (3.6).

(b) Under the general condition A2.1(i), it is sufficient to notice that in (3.5), the terms in $\mu(\cdot)$ have a similar structure as those in $K(\cdot,\cdot)$. By this way, the previous proof can be exactly repeated, using the 'virtual' kernel $K(s,t)=\mu(s)\mu(t)$. ■

### *5.2 Proof of Proposition 3.1*

(a) One may write

$$\sup_{t\in[0,1]} \mathrm{E}\left(X(t)-\widetilde{X}_r(t)\right)^2 = \max_{k=0,\dots,n-1}\ \sup_{t\in[k\delta_n,(k+1)\delta_n]} \mathcal{C}_r(t,t).$$

- *Case* $r_0 \geq 1$.

We apply the results of Lemma 3.1 with the choices $k=\ell$, $\theta_1=\theta_2=\theta\in[0,1]$ :

$$\mathcal{C}_r(t,t) = \sum_{i,j=0}^{r} L_{i,k}(t)L_{j,k}(t)\frac{(ij\delta_n^2)^{r^*}}{r^{2r^*}(r^*-1)!^2}\iint_{[0,1]^2}\Big((1-v)(1-w)\Big)^{r^*-1}$$
$$\Big\{\mathrm{E}\,X^{(r^*)}(k\delta_n+\theta\delta_n v)X^{(r^*)}(k\delta_n+\theta\delta_n w)-\mathrm{E}\,X^{(r^*)}(k\delta_n+\theta\delta_n v)X^{(r^*)}(k\delta_n+\frac{j\delta_n}{r}w)$$
$$-\mathrm{E}\,X^{(r^*)}(k\delta_n+\frac{i\delta_n}{r}v)X^{(r^*)}(k\delta_n+\theta\delta_n w)+\mathrm{E}\,X^{(r^*)}(k\delta_n+\frac{i\delta_n}{r}v)X^{(r^*)}(k\delta_n+\frac{j\delta_n}{r}w)\Big\}\,dv\,dw.$$

(i) Consider first the case $r^*=r_0$ ($r\geq r_0$). Condition A2.1(ii) may be equivalently written as :

$$\mathrm{E}\,X^{(r_0)^2}(s)+\mathrm{E}\,X^{(r_0)^2}(t)-2\mathrm{E}\,X^{(r_0)}(s)X^{(r_0)}(t)\leq C_1^2\,|s-t|^{2\beta_0} \tag{5.1}$$

and we can also decompose the term $\{\cdot\}$ into :

$$\mathrm{E}\,X^{(r_0)}(k\delta_n+\theta\delta_n v)X^{(r_0)}(k\delta_n+\theta\delta_n w)-\frac{1}{2}\mathrm{E}\,X^{(r_0)^2}(k\delta_n+\theta\delta_n v)-\frac{1}{2}\mathrm{E}\,X^{(r_0)^2}(k\delta_n+\theta\delta_n w)$$
$$-\mathrm{E}\,X^{(r_0)}(k\delta_n+\theta\delta_n v)X^{(r_0)}(k\delta_n+\frac{j\delta_n}{r}w)+\frac{1}{2}\mathrm{E}\,X^{(r_0)^2}(k\delta_n+\theta\delta_n v)+\frac{1}{2}\mathrm{E}\,X^{(r_0)^2}(k\delta_n+\frac{j\delta_n}{r}w)$$
$$-\mathrm{E}\,X^{(r_0)}(k\delta_n+\frac{i\delta_n}{r}v)X^{(r_0)}(k\delta_n+\theta\delta_n w)+\frac{1}{2}\mathrm{E}\,X^{(r_0)^2}(k\delta_n+\frac{i\delta_n}{r}v)+\frac{1}{2}\mathrm{E}\,X^{(r_0)^2}(k\delta_n+\theta\delta_n w)$$
$$+\mathrm{E}\,X^{(r_0)}(k\delta_n+\frac{i\delta_n}{r}v)X^{(r_0)}(k\delta_n+\frac{j\delta_n}{r}w)-\frac{1}{2}\mathrm{E}\,X^{(r_0)^2}(k\delta_n+\frac{i\delta_n}{r}v)-\frac{1}{2}\mathrm{E}\,X^{(r_0)^2}(k\delta_n+\frac{j\delta_n}{r}w)$$

which leads to

$$\mathcal{C}_r(t,t)\leq \delta_n^{2(r_0+\beta_0)}\cdot 2\,C_1^2\sum_{i,j=0}^{r}|L_{i,k}(t)L_{j,k}(t)|\,\frac{(ij)^{r_0}}{r^{2r_0}(r_0-1)!^2}\iint_{[0,1]^2}\Big((1-v)(1-w)\Big)^{r_0-1}$$
$$\Big\{\theta^{2\beta_0}\,|v-w|^{2\beta_0}+\left|\theta v-\frac{jw}{r}\right|^{2\beta_0}+\left|\frac{iv}{r}-\theta w\right|^{2\beta_0}+\left|\frac{iv}{r}-\frac{jw}{r}^{2\beta_0}\right|\Big\}\,dv\,dw.$$

Finally, for $t = k\delta_n + \theta\delta_n$,

$$|L_{i,k}(t)| = \prod_{\substack{p=0 \\ p\neq i}}^{r} \left| \frac{t - k\delta_n - p\frac{\delta_n}{r}}{(i-p)\frac{\delta_n}{r}} \right| = \prod_{\substack{p=0 \\ p\neq i}}^{r} \left| \frac{\theta r - p}{i-p} \right| \leq r^r \tag{5.2}$$

where the last upper bound does not depend on $t$ and $k$, so we get the required result :

$$\sup_{t\in[0,1]} \mathcal{C}_r(t,t) \leq A_1^2(r, C_1)\delta_n^{2(r_0+\beta_0)}.$$

(ii) Now if $r^* = r$ (i.e. $r \leq r_0 - 1$), one may write

$$\mathcal{C}_r(t,t) = \delta_n^{2r} \sum_{i,j=0}^{r} L_{i,k}(t)L_{j,k}(t) \iint_{[0,1]^2} \Big((1-v)^{r-1}(1-w)\Big)^{r-1} \frac{(ij)^r}{r^{2r}(r-1)!^2}$$
$$\int_{k\delta_n+\frac{j\delta_n}{r}w}^{k\delta_n+\theta\delta_n w} \int_{k\delta_n+\frac{i\delta_n}{r}v}^{k\delta_n+\theta\delta_n v} K^{(r+1,r+1)}(y,z)\,\mathrm{d}y\,\mathrm{d}z$$

and the result follows from (5.2) and continuity of $K^{(r+1,r+1)}$ :

$$\sup_{t\in[0,1]} \mathcal{C}_r(t,t) \leq A_1^2\Big(r, \sup_{(s,t)\in[0,1]^2} \left|K^{(r+1,r+1)}(s,t)\right|\Big)\delta_n^{2(r_0+1)}.$$

• *Case* $r_0 = 0$. Condition (5.1) together with relation (3.5) ($t = s$, $k = \ell$) yields to the required result (3.7), namely with $r^* = 0$ and $\beta^* = \beta_0$.

(b) Similar to (a) by using the condition A2.1(iii) and Taylor expansion. ■

### *5.3 Proof of Proposition 3.2*

Since $\dot{s}_{k,r} = (\frac{k}{r} + \frac{1}{2r})\delta_n$ and $\frac{k}{r} - \left[\frac{k}{r}\right] = \frac{\alpha}{r}$ with $\alpha = 0, \dots, r-1$, one has $\dot{s}_{k,r} \in \Big[\left[\frac{k}{r}\right]\delta_n, (\left[\frac{k}{r}\right]+1)\delta_n\Big]$. Then $\widetilde{X}_r(\dot{s}_{k,r})$ is a linear combination of variables $\Big\{X(\left[\frac{k}{r}\right]\delta_n + \frac{j}{r}\delta_n),\ j = 0, \dots, r\Big\}$ and no points of $]s_{k,r}, s_{k+1,r}[$ are involved in this interpolation. Consequently, using condition A2.1(iv) with $a = s_{k,r}$ and $b = s_{k+1,r}$, we obtain

$$\begin{aligned} \mathrm{E}\left(X(\dot{s}_{k,r}) - \widetilde{X}_r(\dot{s}_{k,r})\right)^2 &\geq \mathrm{E}\left(X(\dot{s}_{k,r}) - \overline{X}_{s_{k,r},s_{k+1,r}}(\dot{s}_{k,r})\right)^2 \\ &\geq \frac{C_2^2}{(4r)^{2(r_0+\beta_0)}}\delta_n^{2(r_0+\beta_0)}. \end{aligned}$$

■

### 5.4 Proof of Theorem 4.1

We study $P(\widehat{r} \neq r_0)$ for $n$ sufficiently large to ensure that $r_0 \in \{0, \ldots, p_n\}$.

Let us define $\mathcal{N} = \left\{ \forall r \in \{1, \ldots, p_n\} : \delta_n^{-2r} \frac{1}{nr} \sum_{j=0}^{nr-1} (X(\dot{s}_{j,r}) - \widetilde{X}_r(\dot{s}_{j,r}))^2 < h_n \right\}$, the set corresponding to the case of setting $\widehat{r} = l_0$ and let remark that $\mathcal{N}^c = \bigcup_{r=1}^{p_n} \left\{ \delta_n^{-2r} \frac{1}{nr} \sum_{j=0}^{nr-1} (X(\dot{s}_{j,r}) - \widetilde{X}_r(\dot{s}_{j,r}))^2 \geq h_n \right\}$.

We start from $\{\widehat{r} = r_0\} \supset \{\widehat{r} = r_0\} \cap \mathcal{N}^c$ where

- if $r_0 \geq 1$, one has $\{\widehat{r} = r_0\} \cap \mathcal{N}^c = \mathcal{A}(r_0) \cap \mathcal{B}(r_0)$ with

$$\mathcal{A}(r_0) = \left\{ \delta_n^{-2(r_0+1)} \frac{1}{n(r_0+1)} \sum_{j=0}^{n(r_0+1)-1} (X(\dot{s}_{j,r_0+1}) - \widetilde{X}_{r_0+1}(\dot{s}_{j,r_0+1}))^2 \geq h_n \right\}$$
$$\mathcal{B}(r_0) = \bigcap_{r=1}^{r_0} \left\{ \delta_n^{-2r} \frac{1}{nr} \sum_{j=0}^{nr-1} (X(\dot{s}_{j,r}) - \widetilde{X}_r(\dot{s}_{j,r}))^2 < h_n \right\}$$

- if $r_0 = 0$, $\{\widehat{r} = 0\} \cap \mathcal{N}^c = \left\{ n^{-1} \delta_n^{-2} \sum_{j=0}^{n-1} (X(\dot{s}_{j,r}) - \widetilde{X}_1(\dot{s}_{j,r}))^2 \geq h_n \right\} = \mathcal{A}(0)$.

Next for all $r_0 \geq 0$,

$$P(\widehat{r} \neq r_0) \leq P(\widehat{r} \neq r_0 \cup \mathcal{N}) \leq A(r_0) + B(r_0), \tag{5.3}$$

where $B(0) := 0$ and $A(r_0)$ (with $r_0 \geq 0$), $B(r_0)$ (with $r_0 \geq 1$) are respectively defined by

$$A(r_0) := P\left( \delta_n^{-2(r_0+1)} \frac{1}{n(r_0+1)} \sum_{j=0}^{n(r_0+1)-1} (X(\dot{s}_{j,r_0+1}) - \widetilde{X}_{r_0+1}(\dot{s}_{j,r_0+1}))^2 < h_n \right) \tag{5.4}$$

$$B(r_0) := \sum_{r=1}^{r_0} P\left( \delta_n^{-2r} \frac{1}{nr} \sum_{j=0}^{nr-1} (X(\dot{s}_{j,r}) - \widetilde{X}_r(\dot{s}_{j,r}))^2 \geq h_n \right). \tag{5.5}$$

The following two lemmas (whose proofs are postponed to the end of this section) allow us to get exponential upper bounds for $A(r_0)$ and $B(r_0)$.

**Lemma 5.1** *Let us assume the hypothesis A2.1,*

*(i) if the condition A2.2(i) is fulfilled, we have for all* $n \geq N(A_3, r_0, \beta_0)$*,*

$$A(r_0) \leq 3 \exp\Big(-D_1 \min\Big(n, \delta_n^{-(2-2\beta_0)}\Big)\Big);$$

*(ii) under the condition A2.2(ii), we have for all* $n \geq N(A_3, r_0, \beta_0)$

$$A(r_0) \leq 3 \exp\Big(-D_2 \min\Big(\frac{n}{\ln n}, \delta_n^{-(2-2\beta_0)}\Big)\Big)$$

*where* $D_1$*,* $D_2$ *are positive constants.*

**Lemma 5.2** *Let us assume hypothesis A2.1,*

*(i) if the condition A2.2(i) is fulfilled, we have for all* $n \geq N(A_1, r_0, \beta_0)$

$$B(r_0) \leq 3r_0 \exp\Big(-D_1' h_n \delta_n^{-2\beta_0} \min\Big(n, \delta_n^{-(2-2\beta_0)}\Big)\Big)$$

*(ii) under the condition A2.2(ii), we have for all* $n \geq N(A_1, r_0, \beta_0)$

$$B(r_0) \leq 3r_0 \exp\Big(-D_2' h_n \delta_n^{-2\beta_0} \min\Big(\frac{n}{\ln n}, \delta_n^{-(2-2\beta_0)}\Big)\Big)$$

*where* $D_1'$ *and* $D_2'$ *are positive constants and* $h_n \delta_n^{-2\beta_0} \to \infty$ *as* $n \to \infty$*.*

So with (5.3), we may deduce that for $n$ large enough :

(i) under Assumptions A2.1 and A2.2(i),

$$P(\widehat{r} \neq r_0) \leq 3 \exp\Big(-D_1 \min\Big(n, \delta_n^{-(2-2\beta_0)}\Big)\Big) + 3r_0 \exp\Big(-D_1' h_n \delta_n^{-2\beta_0} \min\Big(n, \delta_n^{-(2-2\beta_0)}\Big)\Big), \quad (5.6)$$

(ii) under Assumptions A2.1 and A2.2(ii),

$$P(\widehat{r} \neq r_0) \leq 3(1 + r_0) \exp\Big(-D_2 \min\Big(\frac{n}{\ln n}, \delta_n^{-(2-2\beta_0)}\Big)\Big) + 3r_0 \exp\Big(-D_2' h_n \delta_n^{-2\beta_0} \min\Big(\frac{n}{\ln n}, \delta_n^{-(2-2\beta_0)}\Big)\Big). \quad (5.7)$$

■

**(a) Case $\mathbf{r = 1, \dots, r_0 - 1}$ ($\mathbf{r_0 \geq 2}$).**

We are in position to apply the result of Lemma 3.1 (see also relation (3.6)) in the case $\mu(t) \equiv \mu$. Using the euclidian division of $k$ (respectively $\ell$) by $r$ : $k = p_1 r + q_1$ (resp. $\ell = p_2 r + q_2$), with $p_i \in \{0, \dots, n-1\}$ and $q_i \in \{0, \dots, r-1\}$ for $i = 1, 2$, we may write $\dot{s}_{k,r} = k\frac{\delta_n}{r} + \frac{\delta_n}{2r} = p_1\delta_n + (q_1 + 0.5)\frac{\delta_n}{r}$ and the expression of interest becomes

$$\max_{0\leq k\leq nr-1} \sum_{\ell=0}^{nr-1} |\mathcal{E}_r(k,\ell)| = \max_{p_1,q_1} \sum_{p_2=0}^{n-1} \sum_{q_2=0}^{r-1} |\mathcal{E}_r(p_1 r + q_1, p_2 r + q_2)|$$

First since $r \leq r_0 - 1$, Lemma 3.1 implies

$$\mathcal{E}_r(p_1 r + q_1, p_2 r + q_2) = \sum_{i=0}^{r} \sum_{j=0}^{r} L_{i,p_1}(\dot{s}_{p_1 r+q_1,r}) L_{j,p_2}(\dot{s}_{p_2 r+q_2,r}) \frac{(\frac{i}{r}\delta_n)^r (\frac{j}{r}\delta_n)^r}{(r-1)!^2} \times \int_0^1 \int_0^1 (1-v)^{r-1}(1-w)^{r-1} \mathcal{R}_{\delta_n,i,j}(v,w)\, dv\, dw$$

with

$$\mathcal{R}_{\delta_n,i,j}(v,w) = K^{(r,r)}\Big(p_1\delta_n + (q_1 + \frac{1}{2})\frac{\delta_n}{r}v, p_2\delta_n + (q_2 + \frac{1}{2})\frac{\delta_n}{r}w\Big) + K^{(r,r)}\Big(p_1\delta_n + i\frac{\delta_n}{r}v, p_2\delta_n + j\frac{\delta_n}{r}w\Big) - K^{(r,r)}\Big(p_1\delta_n + i\frac{\delta_n}{r}v, p_2\delta_n + (q_2 + \frac{1}{2})\frac{\delta_n}{r}w\Big) - K^{(r,r)}\Big(p_1\delta_n + (q_1 + \frac{1}{2})\frac{\delta_n}{r}v, p_2\delta_n + j\frac{\delta_n}{r}w\Big)$$

what can be rewritten :

$$\begin{aligned}\mathcal{R}_{\delta_n,i,j} &= \int_{p_2\delta_n + \frac{j\delta_n}{r}w}^{p_2\delta_n+(q_2+1/2)\frac{\delta_n}{r}w} \int_{p_1\delta_n+i\frac{\delta_n}{r}v}^{p_1\delta_n+(q_1+1/2)\frac{\delta_n}{r}v} K^{(r+1,r+1)}(s,t)\, ds\, dt \\ &\leq c\delta_n^2\end{aligned}$$

where, here and in all the following , $c$ denotes a generic positive constant (independent of $n, p_1, p_2$) whose value may vary from line to line. Now, using the independence of $L_{i,p_1}(\dot{s}_{p_1 r+q_1,r})$ and $L_{i,p_2}(\dot{s}_{p_2 r+q_2,r})$ from $p_1$, $p_2$ and $n$, see (5.2), we obtain

$$|\mathcal{E}_r(p_1 r + q_1, p_2 r + q_2)| \leq c\, \delta_n^{2r+2}.$$

Here constant $c$ depends on $q_1$, $q_2$ which are bounded by $r_0$, so we conclude with

$$\max_{k=0,\dots,nr-1} \frac{\delta_n^{-2r}}{n} \sum_{\ell=0}^{nr-1} |\mathcal{E}_r(k,\ell)| \leq C_1\, \delta_n^2.$$

**(b) Case $\mathbf{r = r_0}$ $\mathbf{(r_0 \geq 1)}$ or $\mathbf{r = r_0 + 1}$.**

Now, we write

$$\max_{0\leq k\leq nr-1}\sum_{\ell=0}^{nr-1}|\mathcal{E}_r(k,\ell)| = \max_{p_1,q_1}\sum_{\substack{p_2=0\\|p_2-p_1|\geq 2}}^{n-1}\sum_{q_2=0}^{r-1}|\mathcal{E}_r(p_1r+q_1,p_2r+q_2)|$$
$$+\max_{p_1,q_1}\sum_{p_2=p_1-1}^{p_1+1}\sum_{q_2=0}^{r-1}|\mathcal{E}_r(p_1r+q_1,p_2r+q_2)|$$

This decomposition is compulsory as the conditions A2.2(i) or A2.2(ii) on $K^{(r_0+1,r_0+1)}$ cannot be used on the diagonal.

**Case 1 :** Order of $\mathcal{E}_r(p_1r+q_1,p_2r+q_2)$ when $|p_2-p_1|\geq 2$ and A2.2(i) is fulfilled.

Lemma 3.1 implies

$$\mathcal{E}_r(p_1r+q_1,p_2r+q_2) = \sum_{i=0}^{r}\sum_{j=0}^{r}L_{i,p_1}(\dot{s}_{p_1r+q_1,r})L_{j,p_2}(\dot{s}_{p_2r+q_2,r})\frac{(\frac{i}{r}\delta_n)^{r_0}(\frac{j}{r}\delta_n)^{r_0}}{(r_0-1)!^2}$$
$$\times\int_0^1\int_0^1(1-v)^{r_0-1}(1-w)^{r_0-1}\mathcal{R}_{\delta_n,i,j}(v,w)\;dv\;dw \quad (5.8)$$

with

$$\mathcal{R}_{\delta_n,i,j}(v,w) = K^{(r_0,r_0)}\Big(p_1\delta_n+(q_1+\frac{1}{2})\frac{\delta_n}{r}v,p_2\delta_n+(q_2+\frac{1}{2})\frac{\delta_n}{r}w\Big)$$
$$+K^{(r_0,r_0)}\Big(p_1\delta_n+i\frac{\delta_n}{r}v,p_2\delta_n+j\frac{\delta_n}{r}w\Big)-K^{(r_0,r_0)}\Big(p_1\delta_n+i\frac{\delta_n}{r}v,p_2\delta_n+(q_2+\frac{1}{2})\frac{\delta_n}{r}w\Big)$$
$$-K^{(r_0,r_0)}\Big(p_1\delta_n+(q_1+\frac{1}{2})\frac{\delta_n}{r}v,p_2\delta_n+j\frac{\delta_n}{r}w\Big). \quad (5.9)$$

what can be rewritten (since $|p_2-p_1|\geq 2$)

$$\mathcal{R}_{\delta_n,i,j} = \int_{p_2\delta_n+\frac{j\delta_n}{r}w}^{p_2\delta_n+(q_2+1/2)\frac{\delta_n}{r}w}\Big(\int_{p_1\delta_n+i\frac{\delta_n}{r}v}^{p_1\delta_n+(q_1+1/2)\frac{\delta_n}{r}v}K^{(r_0+1,r_0+1)}(s,t)\;ds\Big)\;dt$$
$$=\int_{\mathcal{S}_{\delta_n,i,j}}K^{(r_0+1,r_0+1)}(s,t)\;ds\;dt, \quad (5.10)$$

Here the strip $\mathcal{S}_{\delta_n,i,j}$ does not contain the diagonal. Moreover the quantity $|\mathcal{S}_{\delta_n,i,j}| = \left(\frac{\delta_n}{r}\right)^2 v|q_1+\frac{1}{2}-i|w|q_2+\frac{1}{2}-j|$ never vanishes and is independent of $p_1$ and $p_2$. So, condition A2.2(i) implies $|\mathcal{R}_{\delta_n,i,j}(v,w)|\leq c\delta_n^2$. Now, using the independence of $L_{i,p_1}(\dot{s}_{p_1r+q_1,r})$ and $L_{i,p_2}(\dot{s}_{p_2r+q_2,r})$ about $p_1$, $p_2$ and $n$, see (5.2), we obtain

$$|\mathcal{E}_r(p_1r+q_1,p_2r+q_2)|\leq c\,\delta_n^{2r_0+2}.$$

Note that here the constant $c$ depends of $q_1$, $q_2$ which are bounded by $r_0$.

Finally,

$$\frac{\delta_n^{-2(r_0+\beta_0)}}{n}\max_{p_1,q_1}\sum_{\substack{p_2=0\\|p_2-p_1|\geq 2}}^{n-1}\sum_{q_2=0}^{r-1}|\mathcal{E}_r(p_1r+q_1,p_2r+q_2)|\leq c\,\delta_n^{2(1-\beta_0)} \tag{5.11}$$

**Case 2 :** Order of $\mathcal{E}_r(p_1r+q_1,p_2r+q_2)$ when $|p_2-p_1|\geq 2$ and A2.2(ii) is fulfilled.

We start from the equations (5.8)-(5.10), for e.g. $p_2\geq p_1+2$, $j\leq q_2$ and $i\leq q_1$. If $(t,s)\in\mathcal{S}_{\delta_n,i,j}$,

$$|s-t|\geq\left|p_1\delta_n+(q_1+\frac{1}{2})\frac{\delta_n}{r}-p_2\delta_n\right|\geq\delta_n(p_2-p_1-1)$$

since $q_1+\frac{1}{2}\leq r$. Note that this bound is the same in the cases $(j\leq q_2,i>q_1)$, $(j>q_2,i\leq q_1)$ or $(j>q_2,i\leq q_1)$. Then

$$|\mathcal{R}_{\delta_n,i,j}|\leq c\delta_n^{2\beta_0}\frac{1}{(p_2-p_1-1)^{2-2\beta_0}}$$

which implies in turn

$$\begin{aligned}\frac{\delta_n^{-2(r_0+\beta_0)}}{n}\sum_{p_2\geq p_1+2}\sum_{q_2=0}^{r-1}|\mathcal{E}_r(p_1r+q_1,p_2r+q_2)| &\leq \frac{c}{n}\sum_{t=1}^{n}\frac{1}{t^{2-2\beta_0}}\\ &\leq \frac{c}{n}\Big(1+\int_1^n\frac{1}{u^{2-2\beta_0}}du\Big)\end{aligned}$$

- If $\beta_0=1/2$, then

$$\frac{\delta_n^{-2(r_0+\beta_0)}}{n}\sum_{p_2\geq p_1+2}\sum_{q_2=0}^{r-1}|\mathcal{E}_r(p_1r+q_1,p_2r+q_2)|\leq c\frac{\ln(n)}{n}.$$

- If $\beta_0>1/2$, $\displaystyle\frac{\delta_n^{-2(r_0+\beta_0)}}{n}\sum_{p_2\geq p_1+2}\sum_{q_2=0}^{r-1}|\mathcal{E}_r(p_1r+q_1,p_2r+q_2)|\leq cn^{2(\beta_0-1)}$,

- If $\beta_0<1/2$, $\displaystyle\frac{\delta_n^{-2(r_0+\beta_0)}}{n}\sum_{p_2\geq p_1+2}\sum_{q_2=0}^{r-1}|\mathcal{E}_r(p_1r+q_1,p_2r+q_2)|\leq c\frac{1}{n}$

We can do the same for $\displaystyle\frac{\delta_n^{-2(r_0+\beta_0)}}{n}\sum_{p_1\geq p_2+2}\sum_{q_2=0}^{r-1}|\mathcal{E}_r(p_1r+q_1,p_2r+q_2)|$, minimizing

$|t-s|$ by $p_1 - p_2 - 1$. Then under the assumption A2.2(ii), we obtain

$$\frac{\delta_n^{-2(r_0+\beta_0)}}{n} \max_{p_1,q_1} \sum_{\substack{p_2=0 \\ |p_2-p_1|\geq 2}}^{n-1} \sum_{q_2=0}^{r-1} |\mathcal{E}_r(p_1 r + q_1, p_2 r + q_2)| \leq c \max\Big(\frac{\ln n}{n}, \delta_n^{2(1-\beta_0)}\Big). \tag{5.12}$$

**Case 3:** Order of $\mathcal{E}_r(p_1 r + q_1, p_2 r + q_2)$ with $p_2 \in \{p_1 - 1, p_1, p_1 + 1\}$.

Let us start with $p_2 = p_1$ : from (5.8) and (5.9), condition A2.1(i) yields to

$$\begin{aligned}|\mathcal{R}_{\delta_n,i,j}(v,w)| \leq \frac{C_1}{r}\delta_n^{2\beta_0}\Big(\Big|q_1 v - q_2 w + \frac{v-w}{2}\Big|^{2\beta_0} + \Big|q_2 w + \frac{w}{2} - iv\Big|^{2\beta_0} \\ + \Big|q_1 v + \frac{v}{2} - jw\Big|^{2\beta_0} + \Big|jw - iv\Big|^{2\beta_0}\Big) =: M(q_1,q_2,v,w,i,j)\delta_n^{2\beta_0}\end{aligned}$$

Now, using the fact that $L_{i,p_1}(\cdot)$ is bounded by $r^r$, see equation (5.2), one may deduce that

$$\begin{aligned}&|\mathcal{E}_r(p_1 r + q_1, p_1 r + q_2)| \\ &\quad \leq \delta_n^{2\beta_0+2r_0} \sum_{i=0}^{r}\sum_{j=0}^{r} |L_{i,p_1}(\dot{s}_{p_1 r+q_1,r}) L_{j,p_1}(\dot{s}_{p_1 r+q_1,r})| \frac{(ij)^{r_0}}{r^{2r_0}(r_0-1)!^2} \\ &\qquad \times \Big|\int_0^1\int_0^1 (1-v)^{r_0-1}(1-w)^{r_0-1} M(q_1,q_2,v,w,i,j)\, dv\, dw\Big| \\ &\quad \leq c\delta_n^{2\beta_0+2r_0}.\end{aligned} \tag{5.13}$$

Finally, the two other terms $\mathcal{E}_r(p_1 r + q_1, (p_1+1)r + q_2)$ and $\mathcal{E}_r(p_1 r + q_1, (p_1 - 1)r + q_2)$ are bounded using the same methodology.

To conclude the proof, one obtains with relations (5.11), (5.13) :

$$\max_{k=0,\dots,nr-1} \frac{\delta_n^{-2(r_0+\beta_0)}}{n} \sum_{\ell=0}^{nr-1} |\mathcal{E}_r(k,\ell)| \leq c \max\left(\delta_n^{2(1-\beta_0)}, \frac{1}{n}\right)$$

under the condition A2.2(i), whereas under condition A2.2(ii), (5.12)-(5.13) yield

$$\max_{k=0,\dots,nr-1} \frac{\delta_n^{-2(r_0+\beta_0)}}{n} \sum_{\ell=0}^{nr-1} |\mathcal{E}_r(k,\ell)| \leq c \max\left(\delta_n^{2(1-\beta_0)}, \frac{\ln n}{n}\right).$$

### *5.6 Proof of Theorem 4.2*

(i) Clear from Corollary 4.1 and Borel-Cantelli lemma.

(ii) If $\widetilde{r} = \max(\widehat{r}, 1)$, one may write

$$
\begin{aligned}
P\Big(\big|X(t) - \widetilde{X}_{\widetilde{r}}(t)\big| \geq \varepsilon_n\Big) \\
&= P\Big(\big|X(t) - \widetilde{X}_{\widetilde{r}}(t)\big| \geq \varepsilon_n, \widehat{r} = r_0\Big) + P\Big(\big|X(t) - \widetilde{X}_{\widetilde{r}}(t)\big| \geq \varepsilon_n, \widehat{r} \neq r_0\Big) \\
&\leq P\Big(\big|X(t) - \widetilde{X}_r(t)\big| \geq \varepsilon_n\Big) + P\big(\widehat{r} \neq r_0\big) \qquad (5.14)
\end{aligned}
$$

where we have set $r = \max(r_0, 1)$. Now since,

$$\big|X(t) - \widetilde{X}_r(t)\big| \leq \big|X(t) - \widetilde{X}_r(t) - \mathrm{E}\,(X(t) - \widetilde{X}_r(t))\big| + \big|\mathrm{E}\,(X(t) - \widetilde{X}_r(t))\big|,$$

we get :

$$
\begin{aligned}
P\Big(\big|X(t) - \widetilde{X}_r(t)\big| \geq \varepsilon_n\Big) \\
&\leq P\Big(\big|X(t) - \widetilde{X}_r(t) - \mathrm{E}\,(X(t) - \widetilde{X}_r(t))\big| \geq \varepsilon_n - \big|\mathrm{E}\,(X(t) - \widetilde{X}_r(t))\big|\Big).
\end{aligned}
$$

The choice $\varepsilon_n = \eta\sqrt{\ln n}\,\delta_n^{r_0+\beta_0}$ and (3.8) in Proposition 3.1 imply :

$$\varepsilon_n - \big|\mathrm{E}\,(X(t) - \widetilde{X}_r(t))\big| \geq \eta\sqrt{\ln n}\,\delta_n^{r_0+\beta_0}\Big(1 - A_2\frac{\delta_n^{1-\beta_0}}{\eta\sqrt{\ln n}}\Big)$$

so that, $\varepsilon_n - \big|\mathrm{E}\,(X(t) - \widetilde{X}_r(t))\big| \geq \eta\sqrt{\ln n}\,\delta_n^{r_0+\beta_0}\alpha_n$, with $\alpha_n \xrightarrow[n\to\infty]{} 1$.

Moreover, $\mathrm{Var}\,\Big(X(t) - \widetilde{X}_r(t)\Big) \leq \mathrm{E}\,\Big(X(t) - \widetilde{X}_r(t)\Big)^2 \leq A_1^2\delta_n^{2(r_0+\beta_0)}$, so we finally have

$$
\begin{aligned}
P\Big(\big|X(t) - \widetilde{X}_r(t) - \mathrm{E}\,(X(t) - \widetilde{X}_r(t))\big| \geq \varepsilon_n - \big|\mathrm{E}\,(X(t) - \widetilde{X}_r(t))\big|\Big) \\
\leq 2\exp\Big(-\frac{\eta^2\alpha_n^2\ln n}{2A_1^2}\Big)
\end{aligned}
$$

as a consequence of the well-known following lemma :

**Lemma 5.3** *If $X \sim \mathcal{N}(0, \sigma^2)$, $\sigma > 0$, then for all $\varepsilon > 0$,*

$$P\big(|X| \geq \varepsilon\big) \leq 2\exp\Big(-\frac{\varepsilon^2}{2\sigma^2}\Big).$$

Next, from (5.14) and Theorem 4.1,

$$\sum_n P\Big(\big|X(t) - \widetilde{X}_{\widetilde{r}}(t)\big| \geq \eta\sqrt{\ln n}\,\delta_n^{r_0+\beta_0}\Big) < \infty$$

for all $\eta > \sqrt{2}\,|A_1|$. Borel-Cantelli lemma implies that almost surely,

$$\limsup_{n\to\infty} \frac{\delta_n^{-(r_0+\beta_0)}}{\sqrt{\ln n}} \left| X(t) - \widetilde{X}_{\widetilde{r}}(t) \right| \leq \sqrt{2}\,|A_1|\,.$$

Now, the required result follows since one has $(\log_a n)^{2(r_0+\beta_0)}(\ln N)^{-\varepsilon} \to 0$ as $n \to \infty$ for all positive $\varepsilon$ and $a \geq 2$. ■

*5.7 Proof of Lemma 5.1*

Let us set $\kappa = r_0 + 1$, we introduce the quantity

$$Z_\kappa := \frac{\delta_n^{-2(r_0+\beta_0)}}{n\kappa} \sum_{j=0}^{n\kappa-1} (X(\dot{s}_{j,\kappa}) - \widetilde{X}_\kappa(\dot{s}_{j,\kappa}))^2. \tag{5.15}$$

From (5.4), we can write

$$\begin{aligned} A(r_0) &= P\left(Z_\kappa < h_n \delta_n^{-2\beta_0+2}\right) \\ &= P\left(E(Z_\kappa) - Z_\kappa > E(Z_\kappa) - h_n \delta_n^{-2\beta_0+2}\right), \end{aligned}$$

where $E(Z_\kappa) = \frac{\delta_n^{-2(r_0+\beta_0)}}{n\kappa} \sum_{j=0}^{n\kappa-1} E(X(\dot{s}_{j,\kappa}) - \widetilde{X}_\kappa(\dot{s}_{j,\kappa}))^2$. Using Propositions 3.1 and 3.2, we can deduce that $E(Z_\kappa)$ is bounded and belongs to the interval $[A_3^2(r_0, C_2), A_1^2(r_0, C_1)]$. Since $h_n$ is chosen such that $h_n \delta_n^{-2\beta_0+2} \to 0$ for all $\beta_0 \in [0,1[$, one has $\eta_n = EZ_\kappa - h_n \delta_n^{-2\beta_0+2} > 0$ for large enough $n$ and $\eta_n \not\to 0$. Thus we need an asymptotic bound for $P(|E(Z_\kappa) - Z_\kappa| > \eta_n)$. We define the vector $Z = (Z^{(0)}, \ldots, Z^{(n\kappa-1)})^\top$ with

$$Z^{(j)} = \frac{\delta_n^{-(r_0+\beta_0)}}{\sqrt{n\kappa}} (X(\dot{s}_{j,\kappa}) - \widetilde{X}_\kappa(\dot{s}_{j,\kappa})). \tag{5.16}$$

One may write :

$$Z_\kappa - \mathrm{E}\, Z_\kappa = \sum_{j=0}^{n\kappa-1} \left\{ (Z^{(j)} - \mathrm{E}\, Z^{(j)})^2 - \mathrm{E}\,(Z^{(j)} - \mathrm{E}\, Z^{(j)})^2 \right\} + 2 \sum_{j=0}^{n\kappa-1} \mathrm{E}\, Z^{(j)} (Z^{(j)} - \mathrm{E}\, Z^{(j)}),$$

so that

$$P(|E(Z_\kappa) - Z_\kappa| > \eta_n) \leq S_1 + S_2 \tag{5.17}$$

with

$$S_1 := P\Big(\Big|\sum_{j=0}^{n\kappa-1}\Big\{(Z^{(j)} - \mathrm{E}\,Z^{(j)})^2 - \mathrm{E}\,(Z^{(j)} - \mathrm{E}\,Z^{(j)})^2\Big\}\Big| \geq \frac{\eta_n}{2}\Big) \tag{5.18}$$

$$S_2 := P\Big(\Big|\sum_{j=0}^{n\kappa-1}(Z^{(j)} - \mathrm{E}\,Z^{(j)})\mathrm{E}\,Z^{(j)}\Big| \geq \frac{\eta_n}{4}\Big). \tag{5.19}$$

**(a) Study of the term $S_1$**

We consider an orthogaussian basis $\{Y_i\}$, $Y_i$ i.i.d. $\mathcal{N}(0,1)$, of the subspace of $L^2$ spanned by $\{Z^{(j)} - \mathrm{E}\,Z^{(j)}\}_{j=0,\ldots,n\kappa-1}$. Let denote the size of the basis by $d_n$. Then we can write $Z^{(j)} - \mathrm{E}\,Z^{(j)} = \sum_{i=1}^{d_n}\mathrm{Cov}\,\Big(Z^{(j)}, Y_i\Big)Y_i := \sum_{i=1}^{d_n} b_{j,i}Y_i$.

Next, if $Y = (Y_1, \ldots, Y_{d_n})^\top$, we obtain

$$\sum_{j=0}^{n\kappa-1}(Z^{(j)} - \mathrm{E}\,Z^{(j)})^2 = \sum_{k,\ell=1}^{d_n} c_{k,\ell}Y_kY_\ell = Y^\top CY$$

where $c_{k,\ell} = \displaystyle\sum_{j=0}^{n\kappa-1}\mathrm{Cov}\,(Z^{(j)}, Y_k)\mathrm{Cov}\,(Z^{(j)}, Y_\ell)$ and $C = B^\top B$, where $C = \Big(c_{k,\ell}\Big)_{k,\ell=1,\ldots,d_n}$ and $B = \Big(b_{k,\ell}\Big)_{k=0,\ldots,n\kappa-1,\,\ell=1,\ldots,d_n}$. The matrix $C$ is real and symmetric, so there exists an orthogonal matrix $P$ such that $\mathrm{diag}(\lambda_1, \ldots, \lambda_{d_n}) = P^\top CP$, where the quantities $\lambda_i$ are the eigenvalues of the matrix $C$. Then we can transform the quadratic form

$$\sum_{j=0}^{n\kappa-1}(Z^{(j)} - \mathrm{E}\,Z^{(j)})^2 = \sum_{j=1}^{d_n}\lambda_j(P^\top Y)_j^2$$

where $(P^\top Y)_j$ denotes the $j$-th component of the $(d_n \times 1)$ vector $P^\top Y$. As $E(P^\top Y)_j^2 = 1$, we arrive at

$$S_1 = P\Big(\Big|\sum_{k=1}^{d_n}\lambda_k\Big((P^\top Y)_k^2 - 1\Big)\Big| \geq \frac{\eta_n}{2}\Big).$$

Now, using the exponential bound of Hanson and Wright (1971), one gets

$$S_1 \leq \exp\Big(-\min\Big(\frac{D_1\eta_n}{2\max_{i=1,\ldots,d_n}\lambda_i}, \frac{D_2\eta_n^2}{4\sum\lambda_i^2}\Big)\Big).$$

For $n$ large enough, easy calculation implies

$$S_1 \leq \exp\left(-\frac{D_3\eta_n}{\max_{i=1,\ldots,d_n}\lambda_i}\right),$$

where $D_3$ is a positive constant. Next, since $B^\top B$ and $BB^\top$ have the same non zero eigenvalues, we can write

$$\max_{i=1,\dots,d_n} \lambda_i \le \max_{0\le k\le n\kappa-1} \sum_{\ell=0}^{n\kappa-1} \left|\mathrm{E}\left(Z^{(k)} - \mathrm{E}\,(Z^{(k)})\right)\left(Z^{(\ell)} - \mathrm{E}\,(Z^{(\ell)})\right)\right|$$

Remark that the zero-mean variables $Z^{(k)} - \mathrm{E}\,Z^{(k)}$ do satisfy conditions of Proposition 4.1 (with $\mu = 0$) as soon as $X$ satisfies A2.1 and A2.2 (see the Example 5). Therefore the condition A2.2(i) and (4.2) yields

$$\max_{0\le k\le n\kappa-1} \sum_{\ell=0}^{n\kappa-1} \left|\mathrm{E}\left(Z^{(k)} - \mathrm{E}\,(Z^{(k)})\right)\left(Z^{(\ell)} - \mathrm{E}\,(Z^{(\ell)})\right)\right| \le c\,\max\left(\frac{1}{n}, \delta_n^{2-2\beta_0}\right),$$

so that,

$$S_1 \le \exp\Big(-\frac{c\,\eta_n}{\max\big(\frac{1}{n}, \delta_n^{2-2\beta_0}\big)}\Big).$$

Therefore, since $h_n\delta_n^{-2\beta_0+2} \to 0$, one gets for all $n$ large enough that $\eta_n \ge A_3^2/2$ and

$$S_1 \le \exp\Big(-c\min\Big(n, \delta_n^{-2(1-\beta_0)}\Big)\Big). \tag{5.20}$$

Now, assuming A2.2(ii), we get with (4.3)

$$\max_{0\le k\le n\kappa-1} \sum_{\ell=0}^{n\kappa-1} \left|\mathrm{E}\left(Z^{(k)} - \mathrm{E}\,(Z^{(k)})\right)\left(Z^{(\ell)} - \mathrm{E}\,(Z^{(\ell)})\right)\right| \le c\max\left(\frac{\ln n}{n}, \delta_n^{2-2\beta_0}\right).$$

which implies in turn,

$$\begin{aligned} S_1 &\le \exp\Big(-\frac{c\,\eta_n}{\max\big(\frac{\ln n}{n}, \delta_n^{2-2\beta_0}\big)}\Big) \\ S_1 &\le \exp\Big(-c\min\Big(\frac{n}{\ln n}, \delta_n^{-2(1-\beta_0)}\Big)\Big) \end{aligned} \tag{5.21}$$

for $n$ large enough.

**(b) Study of the term $S_2$**

Recall that $S_2$, defined in (5.19), is given by :

$$S_2 = P\Big(\Big|\sum_{j=0}^{n\kappa-1}(Z^{(j)} - \mathrm{E}\,Z^{(j)})\mathrm{E}\,Z^{(j)}\Big| \ge \frac{\eta_n}{4}\Big).$$

Since $X$ is a Gaussian process, one may apply Lemma 5.3 to the zero-mean Gaussian random variable $\sum_{j=0}^{n\kappa-1}(Z^{(j)} - \mathrm{E}\,Z^{(j)})\mathrm{E}\,Z^{(j)}$ (see definition of $Z_j$'s in (5.16)). We get

$$S_2 \le 2\exp\Big(-\frac{\eta_n^2}{32\sigma_n^2}\Big)$$

where $\sigma_n^2 := \operatorname{Var} \sum_{j=0}^{n\kappa-1} (Z^{(j)} - \operatorname{E} Z^{(j)}) \operatorname{E} Z^{(j)}$. Moreover, using results (3.7)-(3.8) of Proposition 3.1, one may successively obtain

$$
\begin{aligned}
\sigma_n^2 &= \sum_{j,k=0}^{n\kappa-1} \Big(\operatorname{E} Z^{(j)}\Big)\Big(\operatorname{E} Z^{(k)}\Big) \operatorname{Cov}\Big(Z^{(j)}, Z^{(k)}\Big) \\
&\leq \kappa^2 n^2 \Big(\max_{j=0,\ldots,n\kappa-1} \left|\operatorname{E} Z^{(j)}\right|\Big)^2 \max_{j=0,\ldots,n\kappa-1} \operatorname{E}(Z^{(j)})^2 \\
&\leq A_1^2 A_3^2\, \delta_n^{2(1-\beta_0)}.
\end{aligned}
$$

Finally, for $n$ large enough we get

$$
S_2 \leq 2 \exp\Big(-c\, \delta_n^{-2(1-\beta_0)}\Big). \tag{5.22}
$$

In conclusion, by collecting results from (5.17), (5.20)-(5.22), the claimed result is proved. ■

### *5.8 Proof of Lemma 5.2*

Recall that the term $B(r_0)$, defined in (5.5), occurs only when $r_0 \geq 1$. We introduce the quantity

$$
T_r := \frac{\delta_n^{-2r}}{nr} \sum_{j=0}^{nr-1} (X(\dot{s}_{j,r}) - \widetilde{X}_r(\dot{s}_{j,r}))^2, \quad \text{for } r = 1, \ldots, r_0,
$$

so that $B(r_0) = \sum_{r=1}^{r_0} P\left(T_r - ET_r \geq \nu_n(r)\right)$ where $\nu_n(r) = h_n - \operatorname{E} T_r$.

• Suppose that $r = r_0$, in that case with (5.15), one has $T_{r_0} = \delta_n^{2\beta_0} Z_{r_0}$. Propositions 3.1 and 3.2 lead to

$$
A_3^2(r_0)\delta_n^{2\beta_0} \leq E(T_{r_0}) \leq A_1^2(r_0)\delta_n^{2\beta_0}
$$

Now using the same bounding method as in Lemma 5.1, with $\eta_n$ replaced by $\delta_n^{-2\beta_0}\nu_n(r_0)$, we obtain, under the condition A2.2(i) and for $n$ large enough, that $\nu_n(r_0) \geq 0$ and

$$
\begin{aligned}
P\Big(\Big(T_{r_0} - E(T_{r_0})\Big) \geq \nu_n(r_0)\Big) &\leq 3 \exp\left(-\frac{c\delta_n^{-2\beta_0}\nu_n(r_0)}{\max(\frac{1}{n}, \delta_n^{2-2\beta_0})}\right) \\
&\leq 3 \exp\left(-c\, h_n \delta_n^{-2\beta_0} \min\Big(n, \delta_n^{2\beta_0-2}\Big)\right)
\end{aligned}
$$

where $h_n\delta_n^{-2\beta_0} \to \infty$. Whereas under A2.2(ii), one obtains

$$
P\Big(\Big(T_{r_0} - E(T_{r_0})\Big) \geq \nu_n(r_0)\Big) \leq 3 \exp\left(-c\, h_n \delta_n^{-2\beta_0} \min\Big(\frac{n}{\ln n}, \delta_n^{2\beta_0-2}\Big)\right)
$$

for $n$ large enough.

• Suppose that $r = 1, \dots, r_0 - 1$ (this case occurs only if $r_0 \geq 2$). The propositions 3.1 and 3.2 imply that

$$A_3^2(r)\delta_n^{2(r_0+\beta_0-r)} \leq E(T_r) \leq A_1^2(r)\delta_n^2$$

In this case, $\nu_n(r)\delta_n^{-2} = (h_n - \mathrm{E}\,T(r))\delta_n^{-2} \geq h_n\delta_n^{-2} - A_1^2 \to \infty$. Now, one may repeat proof of the Lemma 5.1 with $r_0 + \beta_0$ replaced by $r+1$, $\eta_n$ by $\nu_n(r)\delta_n^{-2}$ and with the help of (4.1) :

$$P(T_r - E(T_r) \geq \nu_n(r)) \leq 3 \exp\left(-c\, h_n\delta_n^{-2}\right)$$

for $n$ large enough and where $h_n\delta_n^{-2} \to \infty$. Finally, collecting all the results, one obtains the claimed upper bound for $B(r_0)$. We can notice that in both cases, the convergence is faster than for $A(r_0)$. ■

## References


Baxter, G.: 1956, A strong limit theorem for Gaussian processes, *Proc. Amer. Math. Soc.* **7**, 522–525.

Blanke, D. and Bosq, D.: 1997, Accurate rates of density estimators for continuous time processes, *Statist. Probab. Letters* **33**(2), 185–191.

Blanke, D. and Pumo, B.: 2003, Optimal sampling for density estimation in continuous time, *J. Time Ser. Anal.* **24**(1), 1–24.

Bucklew, J. A.: 1985, A note on the prediction error for small time lags into the future, *IEEE Trans. Inform. Theory* **31**(5), 677–679.

Cuzick, J.: 1977, A lower bound for the prediction error of stationary Gaussian processes, *Indiana Univ. Math. J.* **26**(3), 577–584.

Ditlevsen, S. and Sørensen, M.: 2004, Inference for observations of integrated diffusion processes, *Scand. J. Statist.* **31**(3), 417–429.

Gladyshev, E. G.: 1961, A new limit theorem for stochastic processes with Gaussian increments, *Theory Probab. Applic.* **6**(1), 52–61.

Hanson, D. L. and Wright, F. T.: 1971, A bound on tail probabilities for quadratic forms in independent random variables, *Ann. Math. Statist.* **42**(3), 1079–1083.

Istas, J. and Lang, G.: 1997, Quadratic variations and estimation of the local hölder index of a Gaussian process, *Ann. Inst. H. Poincaré, Probab. Statist.* **33**(4), 407–436.

Lasinger, R.: 1993, Integration of covariance kernels and stationarity, *Stochastic Process. Appl.* **45**, 309–318.

Lindgren, G.: 1979, Prediction of level crossings for normal processes containing deterministic components, *Adv. in Appl. Probab.* **11**(1), 93–117.

Müller-Gronbach, T.: 1996, Optimal designs for approximating the path of a stochastic process, *J. Statist. Plann. Inference* **49**(3), 371–385.

Müller-Gronbach, T. and Ritter, K.: 1997, Uniform reconstruction of Gaussian processes, *Stochastic Process. Appl.* **69**(1), 55–70.

Müller-Gronbach, T. and Ritter, K.: 1998, Spatial adaption for predicting random functions, *Ann. Statist.* **26**(6), 2264–2288.

Plaskota, L., Ritter, K. and Wasilkowski, G.: 2002, Average case complexity of weighted approximation and integration over $R_+$, *J. Complexity* **18**(2), 517–544.

Plaskota, L., Ritter, K. and Wasilkowski, G.: 2004, Optimal designs for weighted approximation and integration of stochastic processes on $[0, \infty($, *J. Complexity* **20**(1), 108–131.

Ritter, K.: 2000, *Average-case analysis of numerical problems*, Lecture Notes in Mathematics, 1733, Springer.

Seleznjev, O.: 1996, Large deviations in the piecewise linear approximation of Gaussian processes with stationary increments, *Adv. in Appl. Probab.* **28**(2), 481–499.

Seleznjev, O.: 2000, Spline approximation of random processes and design problems, *J. Statist. Plann. Inference* **84**(1-2), 249–262.

Seleznjev, O. and Buslaev, A.: 1998, Best approximation for classes of random processes, *Technical Report 13, 14 p.*, Univ. Lund Research Report. http://mech.math.msu.su/∼seleznev/bestapp.ps.

Sköld, M. and Hössjer, O.: 1999, On the asymptotic variance of the continuous-time kernel density estimator, *Statist. Probab. Letters* **44**(1), 97–106.

Speckman, P.: 1979, $L_p$ approximation of autoregressive Gaussian processes, *Technical report*, Dept. of Statistics, Univ. of Oregon.